

A FIELD STUDY ABOUT THE IMPACT OF A VR LEARNING UNIT

Thomas Keller, Fabian Hagen and Elke Brucker-Kley
*ZHAW School of Management and Law
St.-Georgen-Platz 2, 8401 Winterthur, Switzerland*

ABSTRACT

This paper examines the learning success of the Virtual Reality course unit "So small! - So big!" compared to conventional teaching. In the form of a field experiment and three random sample tests, the learning success of 67 subjects was examined. In the analysis of the collected data, the differences between the educational levels and the sexes were examined in addition to the overall comparison. In addition, the individual questions of the random sample tests were checked for anomalies. The analysis showed that the two learning units are to be regarded as equivalent. The subjects achieved short and medium-term learning success with both teaching methods. However, no differences could be found between the two teaching methods; they must be regarded as equivalent. The only exception is the difference between the educational levels in terms of medium-term learning success.

KEYWORDS

Virtual Reality, Education, Field Experiment

1. INTRODUCTION

The possibilities and benefits of the use of new media in teaching at elementary schools have been discussed for some time. In the context of computer use in schools, the terms media pedagogy and media didactics are often used (cf. (Krauthausen 2012), p. 1). In the context of the present work, however, "new media" primarily means the use of so-called learning or exercise programs. With the use of IT resources, specific knowledge can be imparted in many subjects. New teaching concepts are possible. Especially VR and augmented reality (AR) bring completely new possibilities into the classroom. Simple examples are Google Expeditions and the Google Cardboard. A study by Samsung in Germany in 2017 shows: "More than three-quarters of teachers (79%) agree with the statement that thanks to VR, students have the opportunity to make experiences that they would never otherwise have. [...] In addition, the majority of teachers believe that the use of VR in class can increase the motivation of students (74%) and improve their learning success (62%). According to the teachers (58%), the use of the technology can also help in understanding learning concepts. The greatest potential benefits of VR are attributed to the subjects geography/geography (80%), history (74%) and science (62%)".

Interestingly, however, digital media are only used hesitantly in elementary schools. There may be many reasons for this. On the one hand, on the side of teachers, on the other hand in pedagogy and in very new media such as virtual reality, the lack of pedagogical content for use in the classroom.

What Confucius (551-479 B.C.) already knew still applies to today's school teaching: "Tell me, and I will forget. Show me, and I may remember. Involve me, and I will understand." This guideline is not only perfectly suited to the topic of virtual reality, it is also often quoted in connection with competence-oriented learning in the context of Lehrplan 21 (d-edk 2016) (Swiss curriculum for primary and secondary schools), where the focus is on the actions of learners. The aim of this work is therefore to develop and implement concrete learning units for mathematical learning in VR that exhibit a degree of interaction that comes close to doing in the sense of Confucius.

(Keller and Brucker-Kley 2018a) discuss an immersive VR learning unit and its usability amongst pupils with special needs. This learning unit is now used for further investigations regarding the impact on the learning success. The learning unit itself is not introduced in this paper. Please refer to (Keller and Brucker-Kley 2018b) for that purpose.

2. STATE-OF-THE-ART

In the following, important principles and mathematics didactics are examined in order to subsequently deal with immersive learning in the context of VR.

2.1 Didactics of Mathematics

((Krauthausen 2012), p. 3 f.) criticises the fact that computer-assisted learning programs in mathematics that are well-known and widespread on the market today pay too little attention to the current state of research and knowledge in mathematics didactics. The focus is too much on technology instead of content (the subject-specific content) and the programs thus contribute much to the media competence of learners but little to mathematics skills.

((Burrill 2017), p. 316) mentions mathematical accuracy (fidelity) and user experience as central mathematics didactic principles. The mathematical accuracy means that the software should always be mathematically correct, the user experience should not hinder the work with the mathematical task and should promote mathematical thinking.

Learners need to be able to make decisions to expand their thinking. This possibility is also closely related to the complexity of a task, which does not necessarily require complex mathematical requirements (cf. (Geiger 2017), p. 289). According to ((Joubert 2017), p. 20 ff.), while working on a mathematical task, students use means from the so-called "Modes of Production". These include acting (usually in the sense of indicating a solution), formulating (developing hypotheses, solution strategies, etc.) and validating (checking based on evidence, theorems or explanations). Formulation and validation almost always have to be initiated by the teacher.

For the motivation of learners, mathematics itself should increasingly be presented in a way that makes them exciting and captivating themselves, rather than motivating them with other means in the learning programs (cf. (Krauthausen 2012), p. 20): "Effective learning processes are characterized by a high degree of motivation and joy, which however arise from the matter and not from its packaging" ((Krauthausen 2012), p. 21).

Mathematical learning is an important pedagogical task. Mathematical tasks are intended to encourage learners to do something mathematical and thus experience mathematics in the broadest sense (cf. (Joubert 2017), p. 4). All tasks should always contain pragmatic and epistemological aspects. The epistemological aspects refer to the insights to be conveyed to learners while working on a task (cf. (Sinclair and Zazkis 2017), p. 177), whereby the pragmatic value of a task is almost always equated with solving the task (cf. (Sinclair and Zazkis 2017), p. 190). ((Laborde 2011), p. 82) supplements cognitive aspects (what kind of learning the task triggers in the learner at the current state of knowledge), didactic aspects (with what means the task is set) and instrumental aspects (which instruments the learner needs to solve the task).

2.2 Immersive Learning

Learning in virtual worlds is often called "immersive learning". According to ((Höntzsch et al. 2013), p. 3), immersion describes the degree to which individuals perceive that they interact more with their virtual than with their real environment (individual sense of being there). In a virtual reality, immersion seems to be determined by the degree of representation of the learners, their presence and their possibilities of interaction.

((Höntzsch et al. 2013), p. 3) describe with reference to ((Burdea and Coiffet 2003), p. 3) the three "I" of learning with virtual realities: Imagination, immersion and interaction. "Imagination describes the imaginative power of learners to put themselves in the position of a simulation. Real-time visualizations and reactions of the system provide users with immediate feedback on their inputs (interaction). The information is also recorded multimodally [...], i.e. with several senses. This creates a feeling of immersion, i.e. of being

directly involved in the simulated world" ((Höntzsch et al. 2013), p. 3). Imagination, immersion and interaction seem to be important factors for immersive learning.

With regard to pedagogy, ((Geiger 2017), p. 288 f.) points out in the context of mathematics how eminently important it is to select, adapt and implement the tasks in the learning environments. In this context, he points out the importance of cooperation between teachers and researchers in order to anchor well-designed tasks with pedagogically correct approaches in the learning environments and thus improve learning.

A sufficient degree of challenge is important for the learning process (cf. (Geiger 2017), p. 289). ((Höntzsch et al. 2013), p. 4) list the following measures as necessary to avoid overburdening learners in immersive learning environments:

- clear learning objectives, work orders and instructions,
- permanently available background information,
- hints and exercises that stimulate reflection (for example, setting a specific state of the simulation)

((Höntzsch et al. 2013), p. 3) list the possible support of immersion on learning processes in connection with flow and presence experience as a thesis when using three-dimensional virtual worlds. However, with reference to Grunewald (2009), they point out that these effects are also mentioned in the context of computer game addiction. ((Chen 2016), p. 644) also shows a positive effect on the learning (in the context of language learning) of such environments.

According to ((Höntzsch et al. 2013), p. 4), discovering learning leads to an expansion of personal experience space and to the generation and examination of hypotheses, since knowledge in these learning environments is not predetermined but explorative. It is pointed out that the learning environments must be simulated as truthfully as possible in order for the findings to be successfully transferred into reality. It also describes advantages in the depth of information processing, learning success and motivation, the latter not per se resulting in a higher quality or quantity of cognitive processing and skill acquisition.

3. RESEARCH OBJECTIVE

The main objective of this work is to compare the learning success of learning units with VR against those without VR support. In addition, the medium-term learning success is to be researched by making a new survey one month later. Another goal is the question whether there are differences in educational level and gender. In addition, the respondents' opinion of this teaching method is to be questioned and the answers compared with the data collected.

3.1 Research Design – Field Experiment

In this field experiment a slightly modified version of the Pretest-Posttest Control Group Design according to (Campbell and Stanley 1967) is applied. The design provides that a test is performed at the beginning of the field experiment, then the independent variable is changed, and then a test is performed again. In addition to this second test, a new test is performed a month later in this study to measure medium-term learning success.

This research design has a high internal validity, as the pre-test guarantees the equivalence of the test persons. The external validity, on the other hand, is weakened by the fact that it cannot be ruled out that the pre-test influences the subjects in their subsequent behaviour (Campbell & Stanley, 1963, p. 13ff). It could be the case that the pupils then exert themselves more in the following lessons, which would perhaps not have been the case without the pre-test. In the case of the present work, this is attempted to prevent by the students not knowing that they are part of a field experiment.

Four classes from a secondary school in Switzerland serve as comparison groups. These are two classes of level A and two classes of level B. One of the two classes will be in the experimental group (VR unit) and the other one in the control group (regular unit). In total, the comparison groups comprise 87 subjects, 45 of whom were at secondary level A and 42 at secondary level B. The comparison groups are divided into two groups: the first group (VR unit) and the second group (regular unit). Due to failures during the field experiment, 67 subjects were included in the final analysis.

An important factor in the control of interfering variables is the randomisation procedure. The aim is to evenly distribute known and unknown personal factors in order to create a level playing field. In the case of this study, randomization is determined by the class distribution of the schools. Thus, the field experiment must be regarded as quasi-randomized.

A protocol is kept during the execution of the field experiment, which records special features such as assistance or technical problems and contains a short questionnaire for the test persons. The questions will be asked and answered verbally immediately after the course unit has been completed.

4. RESULTS

The following subsections present the findings from the field experiment. A more in-depth discussion of the results can be found here¹.

4.1 Differences Between Experimental and Control Group

The immersive as well as the conventional teaching method have resulted in short and medium-term learning success for the test persons of both comparison groups. The subjects in the experimental group between Test 1 and Test 2 achieved an average learning success of 1.38 points (19.14%). In the control group, the average learning success between the same tests was 1.30 points (18.49%). The medium-term learning success, which is measured by the difference between Test 1 and Test 3, averages 0.64 points (8.88%) for the subjects in the test group. The subjects in the control group recorded an average learning success of 0.45 points (6.40%) between the same tests. The correlation is shown in Figure 1. Average test results of experimental and control group.

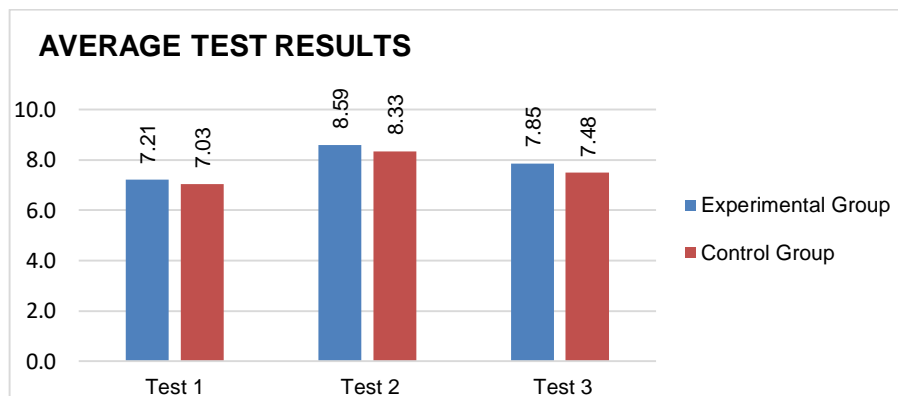


Figure 1. Average test results of experimental and control group

If the average learning outcomes of the subjects in the experimental and control groups are compared, the difference between Test 1 and Test 2 is 0.08 points. The statistical significance is 58.24%, which means that the two teaching methods do not differ. Similar observations can be made for the average learning outcomes between Test 1 and Test 3. There it is a difference of 0.19 points with a statistical significance of 64.71%. These values are illustrated in Table 1.

Table 1. Comparison between experimental with control group

Comparison of Comparative Groups	Experimental Group	Control Group	Difference	Statistical Significance
Learning Success Test 1 to Test 2	1.38	1.30	0.08	58.24%
Learning Success Test 1 to Test 3	0.64	0.45	0.19	64.71%

¹ <https://drive.switch.ch/index.php/s/OELriQqJx8mCver>

In summary, this means that both immersive and conventional teaching methods have resulted in statistically verifiable learning success for the test persons. If the learning outcomes of the different groups are compared, however, there are no differences. The two teaching methods are to be regarded as equivalent in terms of both short-term and medium-term learning success. However, there are tendencies, if the questions of the random sample tests are considered individually, that the tasks with reference to the linear masses were better solved by the test persons of the experimental group and the tasks with reference to the hollow masses were better solved by the test persons of the control group.

4.2 Difference Between Education Levels

In order to be able to make more precise statements about the learning success of the test persons, the test persons of the comparison groups were additionally divided into secondary level A (Sec A) and secondary level B (Sec B). Their average test results are shown in Figure 2. As can be observed, there are no striking differences in the learning outcomes of the test subjects between Test 1 and Test 2. However, if one looks at Test 3, it is noticeable that the test subjects in Test Group Sec A achieved the highest average score, which was not apparent before.

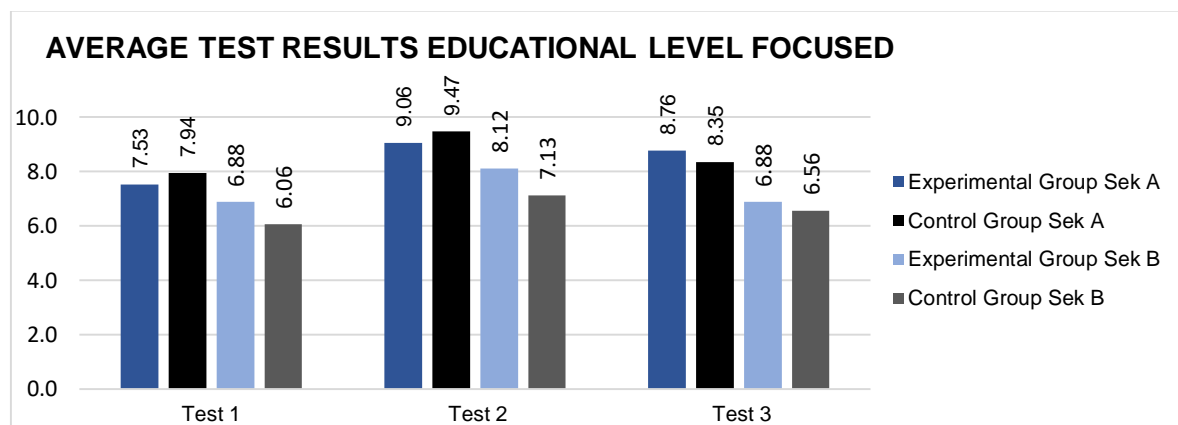


Figure 2. Average test results of different educational levels

The average learning success between Test 1 and Test 3 was 1.53 points (16.33%) for the subjects in the test group's Sec A group. The average learning success achieved by the test subjects in the control group's Sec A was 0.41 points (5.16%), which is considerably lower. If the average learning successes between Test 1 and Test 3 of the subjects in the Sec A of both comparison groups are compared, this leads to the list in Table 1. The difference between the average learning successes is 0.83 points. The statistical significance is 88.50%. This means that the difference is not statistically relevant, but there is a tendency that should be further investigated.

Table 2. Learning success between test 1 and 2

Comparison of Subjects of Sek A	Experimental Group	Control Group	Difference	Statistical Significance
Learning Success Test 1 to Test 3	1.24	0.41	0.83	88.50%

If the average learning success of the subjects in Sec A is compared with that of the subjects in Sec B of the experimental group between Test 1 and Test 3, this leads to a difference of 1.24 points. In addition to the high learning success of the subjects in Sec A, this is mainly due to the fact that the subjects in Sec B did not achieve any learning success in these tests. With a statistical significance of 99.01%, this difference is considered statistically relevant. This correlation can be seen in Table 3.

Table 3. Comparison of learning success between levels

Comparison of Experimental Group	Subjects of Sek A	Subjects of Sek B	Difference	Statistical Significance
Learning Success Test 1 to Test 3	1.24	0.00	1.24	99.01%

In summary, this means that immersive teaching methods led to verifiable short and medium-term learning success in subjects in Sec A. This contrasts with the subjects in Sec B, who only achieved short-term learning success. There is therefore a difference between the educational levels with regard to medium-term learning success with VR learning units.

4.3 Gender Differences

A further possibility of classification is the subdivision by gender. As can be seen in Figure 3, the average test results are similar. The female test subjects in the test group who achieved the highest learning success between Test 1 and Test 2 with 1.77 points (27.06%) are notable. This finding is particularly evident when the focus is placed on the individual questions of the random sample tests. For example, the female respondents in the test group in question 1a, which has a strong relation to the linear measures in the learning unit, have by far the highest average learning success with 225.00% between test 1 and test 2.

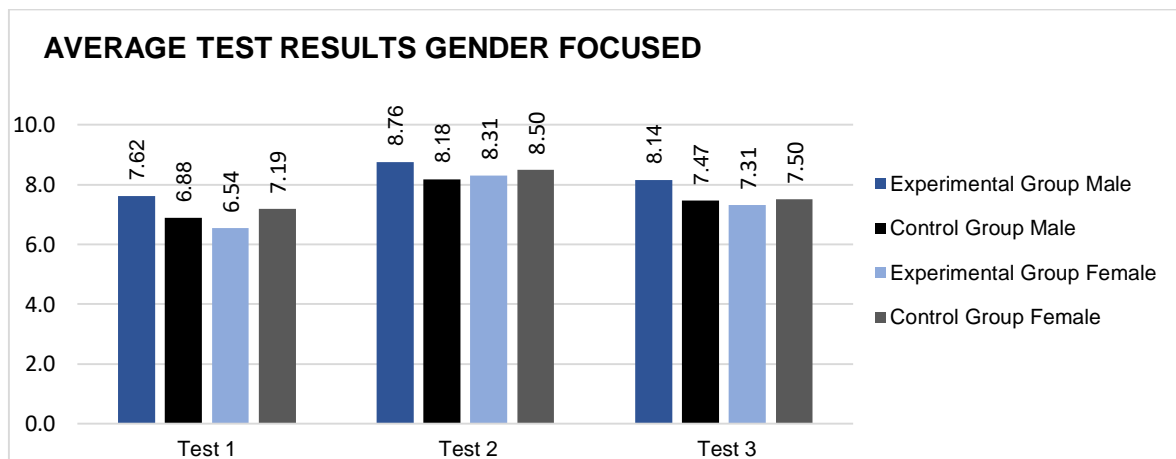


Figure 3. Average test results of genders

As Figure 4 shows, the female respondents in the control group are far below this with an average learning success of 128.57%. The male subjects in the experimental group were also lower with an average learning success of 66.67%. Similar observations can be made for learning success between Test 1 and Test 3. With 125.00% learning success, the female subjects in the experimental group are far ahead of the female subjects in the control group with 57.14% and the male subjects in the experimental group with 66.67%.

The tendency of the female subjects in the experimental group to show high learning progress can also be applied to most other questions, albeit to a lesser extent. In particular, it is the tasks related to the linear measures in which the subjects in the experimental group show greater learning success than the control group.

If, on the other hand, the average learning successes between Test 1 and Test 2 of the female subjects in the test and control groups are compared, the difference is 0.46 points, with the test group being higher. The statistical significance is 78.39%. A similar observation can be made between Test 1 and Test 3, where the difference is also 0.46 points. Here the statistical significance is somewhat lower at 74.82%. This means that both learning units are of equal value for the female subjects.

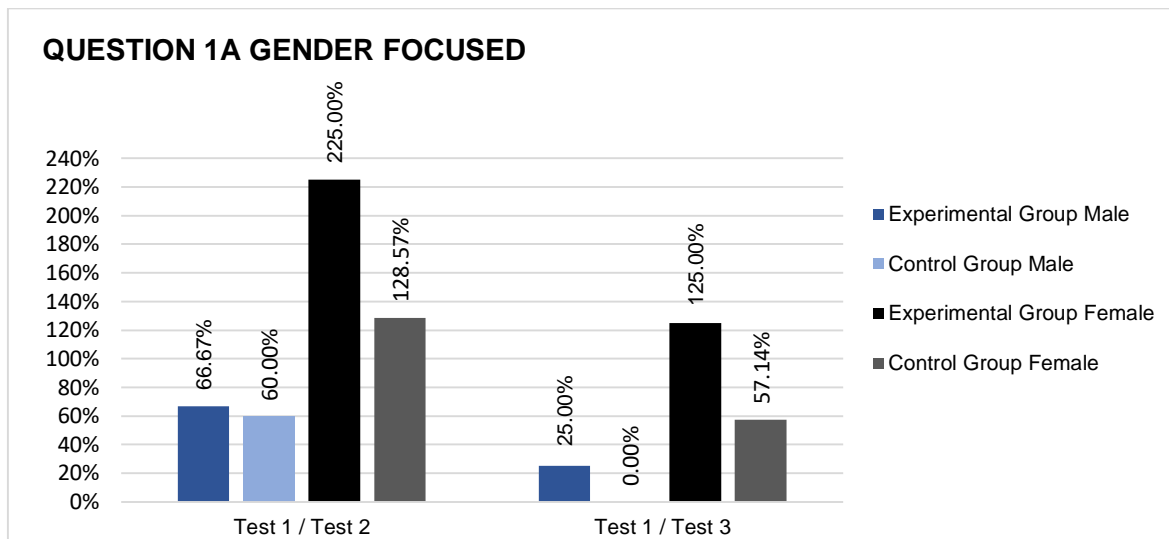


Figure 4. Question 1A visualized according gender

If the average learning successes of the male and female subjects in the comparison group are compared, there is a difference of 0.63 points, with the female subjects being higher.

In summary, this means that there are no differences between the subjects in the experimental and control groups when the focus is on the sexes. The VR and conventional learning units are to be regarded as equivalent. In addition, there are no differences between male and female subjects who used the VR unit. However, there is a tendency that the female subjects benefited more from the VR units than the male subjects.

5. CONCLUSION

The analysis of the data showed that both learning units resulted in short and medium-term learning success for the test persons, but that there were no differences between the two teaching methods. A difference can be observed between the educational levels. Thus, in contrast to the subjects at Sec B, the subjects at Sec A had better learning success in the medium term. Statistically, there are no significant differences between the sexes. However, there are indications that female subjects benefit more than male subjects. This is a finding that would make it possible to provide targeted support for girls, especially in technically related fields.

The above results revealed trends that could not be answered in the present field experiment. They are described in the following and hypotheses for future research are derived from them.

During the data analysis, there were more and more indications that female subjects had greater learning success with the VR course unit than with the conventional course unit. It was also found that they tended to achieve greater learning success than male subjects who did the VR unit. However, this tendency could not be statistically proven in the context of the present study. It concerns both short and medium-term learning success.

The effectiveness of the VR learning unit differs between linear measures and hollow measures. Linear measures performed better than hollow measures. The assumption suggests that the design criteria of the virtual space play a role here.

Furthermore, an important limitation must also be referred to. It cannot be estimated what influence the effect of the new had on the test persons. This can only be avoided if the test persons work with virtual reality over a longer period of time and this technology thus loses the attraction of the new.

Finally, it must be noted that the use of VR learning units does not per se lead to better learning success. As with conventional teaching methods, the quality and design of such VR learning units plays a decisive role.

REFERENCES

- Burdea, Grigore C, and Philippe Coiffet 2003 *Virtual Reality Technology*. 2nd Edition. New Jersey: John Wiley & Sons.
- Burrill, Gail 2017 Designing Interactive Dynamic Technology Activities to Support the Development of Conceptual Understanding. *In* *Digital Technologies in Designing Mathematics Education Tasks: Potential and Pitfalls*. Allen Leung and Anna Baccaglioni-Frank, eds. Pp. 303–328. Mathematics Education in the Digital Era. Switzerland: Springer.
- Campbell, Donald Thomas, and Julian Cecil Stanley 1967 *Experimental and Quasi-Experimental Designs for Research*. 2. print. Boston: Houghton Mifflin Comp.
- Chen, Yu-Li 2016 The Effects of Virtual Reality Learning Environment on Student Cognitive and Linguistic Development. *The Asia-Pacific Education Researcher* 25(4): 637–646.
- d-edk 2016 Lehrplan 21: Gesamtausgabe. InternetDocument. https://v-ef.lehrplan.ch/lehrplan_printout.php?e=1&k=1.
- Geiger, Vince 2017 Designing for Mathematical Applications and Modelling Tasks in Technology Rich Environments. *In* *Digital Technologies in Designing Mathematics Education Tasks: Potential and Pitfalls*. Allen Leung and Anna Baccaglioni-Frank, eds. Pp. 285–302. Mathematics Education in the Digital Era. Switzerland: Springer.
- Höntzsch, Susan, Uwe Katzky, Klaus Bredl, Frank Kappe, and Dirk Krause 2013 *Simulationen Und Simulierte Welten: Lernen in Immersiven Lernumgebungen*. InternetDocument. <http://13t.tugraz.at/index.php/LehrbuchEbner10/article/download/102/108>.
- Joubert, Marie 2017 Revisiting Theory for the Design of Tasks: Special Considerations for Digital Environments. *In* *Digital Technologies in Designing Mathematics Education Tasks: Potential and Pitfalls*. Allen Leung and Anna Baccaglioni-Frank, eds. Pp. 17–40. Mathematics Education in the Digital Era. Switzerland: Springer.
- Keller, Thomas, and Andreas Hebeisen and Elke Brucker-Kley 2018a Integration of Children with Special Needs in Mathematics Through Virtual Reality. *In* Pp. 30–37. <http://www.iadisportal.org/digital-library/integration-of-children-with-special-needs-in-mathematics-through-virtual-reality>, accessed May 22, 2019.
- 2018b Integration of Children with Special Needs in Mathematics through Virtual Reality. *In* Pp. 30–37. <http://www.iadisportal.org/digital-library/integration-of-children-with-special-needs-in-mathematics-through-virtual-reality>, accessed February 22, 2019.
- Krauthausen, Günter 2012 *Digitale Medien Im Mathematikunterricht Der Grundschule*. Heidelberg: Spektrum Akademischer Verlag.
- Laborde, Colette 2011 Designing Substantial Tasks to Utilize ICT in Mathematics Lessons. *In* *Mathematics Education with Digital Technology*. Adrian Oldknow and Carol Knights, eds. Pp. 75–83. New York: Continuum International Publishing Group.
- Samsung Studie: Lehrer Sehen Großes Potenzial Für Die Nutzung von Virtual Reality Im Unterricht 2017 PressRelease. Schwalbach/Ts. https://news.samsung.com/de/wp-content/themes/sw_newsroom/download.php?id=NfO684PPbcmccJ0mMOMjsg%3D%3D.
- Sinclair, Nathalie, and Rina Zazkis 2017 Everybody Counts: Designing Tasks for TouchCounts. *In* *Digital Technologies in Designing Mathematics Education Tasks: Potential and Pitfalls*. Allen Leung and Anna Baccaglioni-Frank, eds. Pp. 175–192. Mathematics Education in the Digital Era. Switzerland: Springer.